DOCUMENT RESUME

ED 345 511 FL 020 174

AUTHOR Young, Richard; Milanovic, Michael
TITLE Discourse Variation in Oral Proficiency Interviews.
PUB DATE Feb 92
NOTE 37p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Discourse Analysis; *English (Second Language); Interpersonal Communication; *Interviews; *Language Proficiency; *Language Role; *Language Styles; *Language Tests; Second Language Learning; Verbal Tests
IDENTIFIERS *Oral Proficiency Testing

ABSTRACT

A study of the discourse structure of oral language proficiency interviews focused on (1) one principal discourse variable, topic, for analyzing contingency and goal orientation in dyadic interactions, and (2) contextual factors (interlocutor, theme, task, participant gender). Data came from 30 dyadic oral interviews in English as a Second Language recorded in Brazil and Italy. The interview was part of the First Certificate in English examination of the University of Cambridge (England) Local Examinations Syndicate. Portions of the interview analyzed included three tasks: discussion based on photographs; relation of a printed passage to the photographs; and expression of personal preferences about items in a list of activities related to the interview's theme. The study was principally exploratory and descriptive. Results are discussed in terms of the characteristics of native-speaker/non-native-speaker oral interaction. It was found that the two parties made very different contributions to the discourse, with the examiner exerting a controlling influence and the examinee having a more reactive role. Contextual factors found to affect only candidate discourse included individual differences among examiners, especially gender, and task. Contextual influences on the examiner's goal orientation appeared to include gender and interview theme. The major influence on discourse as a whole was task. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED345511

DISCOURSE VARIATION IN ORAL PROFICIENCY INTERVIEWS

Richard Young
Southern Illinois University

Michael Milanovic
University of Cambridge Local Examinations Syndicate

(To appear in *Studies in Second Language Acquisition, 14*(4), December 1992)

Please address correspondence to:

Richard Young
Department of Linguistics
Southern Illinois University
Carbondale, IL 62901-4517

Phone: (618) 453–3411
FAX: (618) 453–6527
Bitnet: GA3980 at SIUCVMB

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Richard Young
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 020 174

BEST COPY AVAILABLE

In this paper a theoretical model of dyadic NS-NNS discourse is proposed in which discourse is described in terms of three features: interactional contingency, the goal orientation of participants, and dominance. The model is then used to study the discourse of 30 dyadic oral interviews of the Cambridge First Certificate in English examination. Results demonstrate the effectiveness of the model in abstracting the structure of oral interview discourse. They show that the discourse of oral proficiency interviews is characterized by greater reactiveness by NNS candidates and greater orientation toward goals by NS examiners. Variation in the structure of the discourse is also investigated in this study. This is shown to be related to the examiner, the theme of the interview, the task in which the participants are engaged, and the gender of examiner and candidate.

INTRODUCTION

In recent years much effort has gone into the study of the speech of second language learners in different social contexts and into relating the variation in interlanguage phonology, morphology, and syntax observed in learners' speech to certain features of context. It has been established that much of the variation in learners' spoken language is rule governed and relates in a systematic way to contextual factors, including the particular task in which a learner is engaged, the interlocutor or interlocutors with whom a learner is interacting, the topic or theme that the learner is talking about, and the linguistic environment of the variable form.[1]

These studies of phonological, morphological, and syntactic variation in interlanguage speech lead us to expect that features of spoken interlanguage discourse may be similarly sensitive to the context in which the discourse is constructed and pioneering work in describing variation in interlanguage discourse has indicated that this is, indeed, the case (Takahashi, 1989; Woken & Swales, 1989; Zuengler, 1989a, b). The present study contributes to this developing research tradition with an investigation of the possible influence on the discourse of oral proficiency interviews[2] of the four variables of interlocutor, theme, task, and gender.

In addition to the insights that it may bring into the structure of interlanguage discourse, the investigation of variation in oral proficiency interviews is also an important practical undertaking. As Douglas and Selinker (1985) and Tarone (1988) have pointed out, the abundant evidence linking interlanguage variation to features of the context in which it is elicited has important implications for the design of instruments that attempt to test and evaluate learner's proficiency in a second language. In evaluating learners' proficiency, developers and users of language tests in effect make generalizations from a sample of interlanguage elicited in one context to learners' performance in other and different contexts. What we now know about the systematic influence of context on learner's language may give us pause before we generalize from learners' performance elicited in one context to performance in others.

PREVIOUS RESEARCH ON DISCOURSE IN DYADIC INTERACTIONS

We propose to concentrate in this paper on describing three aspects of the discourse of oral proficiency discourse: interactional contingency, goal orientation, and dominance. These three aspects have been discussed at length in recent research on interactions between native and nonnative speakers of English (van Lier, 1989; Zuengler 1989a, b). We choose these three aspects because they appear to be the most promising in elucidating the discourse structure of oral proficiency interviews and because they may help us to make a principled comparison between oral interviews and the discourse created in other forms of NS-NNS interactions.

Interactional Contingency

The discourse structure of oral proficiency interviews has been discussed in detail by van Lier (1989), who underscores the differences between the discourse of oral interviews and conversation. From the theoretical perspective of conversation analysis, van Lier proposes a model of dyadic interaction based on the work of Jones and Gerard (1967). This model describes dyadic interaction in terms of two dimensions: contingency and goal orientation. Contingency is the social determinant of the structure of interactions: that is, how participants react to each other, while goal orientation reflects the internal goals of each speaker separately.[3] The model promises to help in describing and explaining the dynamic structure of dyadic discourse since contingency and goal orientation may be seen as ways of describing local and global discourse dynamics. Contingency is a way of looking at the local dynamic of how participants create shared meanings in any one exchange, while goal orientation focuses on how those meanings evolve and change over longer stretches of discourse.

We have narrowed the social-psychological model of Jones and Gerard by defining contingency as a property of sequences of speech between two interactants. In our version of the model, a contingent utterance is one in which the content and often the form of the utterance depend in some way on a previous utterance. When the utterance is contingent upon a previous utterance by the other participant, we describe the relationship between the two utterances as "reactiveness." The most obvious examples of reactively contingent utterances are adjacency pairs (Schegloff, 1978) such as question-answer sequences and topic continuity across conversational turns. However, contingency may also exist when conversational participants search for a mutually agreed meaning as in negotiation of meaning sequences (Hatch, 1983) or when superficially different topics are linked across turns by conversational implicature (Grice, 1989). Contingency is thus a property of adjacent turns in dialogue in which the topic of the preceding turn is coreferential with the topic of the following turn. In the two special cases of negotiation of meaning and conversational implicature, coreference of topics in adjacent turns is established by recourse to an underlying and unstated proposition linking the two.

Goal Orientation

The other dimension of the Jones and Gerard model represents the speakers' attempt to realize certain internal goals or plans through the interaction. As Jones and Gerard (1967) point out, "although normally these plans are vague and implicit, ... a plan tends to become quite prominent in awareness when important goals are at stake" (p. 506). In the context of linguistic interaction, goal orientation is evident in the discourse when topics persist over a large number of turns, or when a speaker returns to a topic raised earlier after a number of intervening turns on a different topic. Goal orientation is particularly evident in dyadic interactions in cases when one participant is following an agenda established in advance. In the oral proficiency interviews which form the object of the present study, this is evident when the native speaking examiner follows the procedure established for such interviews by the examination board. On the other hand, lack of goal orientation is evident in the rapid decay of topics over one or two turns.

Jones and Gerard combine the two dimensions of contingency and goal orientation in four styles of dyadic interaction. *Pseudocontingency* is characterized by a high degree of goal orientation but little reactiveness, *asymmetrical contingency* is characterized by a high degree of goal orientation by one party and a high degree of reactiveness by the other party, *reactive contingency* is characterized by little goal orientation by either party but a high degree of reactiveness by both, and *mutual contingency* is identified as a high degree of goal orientation and reactiveness by both parties. These four styles are represented graphically in Figure 1, in which solid arrows represent the predominant or major source of influence on each turn (R) by participant A or B and unfilled arrows represent the less important or minor source. The figure represents the flow of time from top to bottom, with A's turns on the left and B's turns on the right.

Figure 1 about here

In conducting the present study we have operationalized and refined the Jones and Gerard model with the aim of increasing our understanding of dyadic linguistic interaction in general and the discourse which is constructed between native and nonnative speakers of English in oral proficiency interviews in particular.

Dominance

Absent from the work of Jones and Gerard is the variable of dominance, which has been shown by SLA researchers to be an important dimension of dyadic interaction. Dominance in dyadic conversational interaction may be defined as the tendency for one participant to control the discourse by various means, including holding the floor, initiating and terminating topics, and controlling the other participant's access to the floor by means of interruptions and questions. Conversational dominance by one participant in a dyad limits the other participant's right to speak. In research on NS-NNS interactions, dominance by either party has been shown to be a function of the linguistic proficiency of the NNS (Woken & Swales, 1989), the relative expertise of the two participants in the subject-matter of the conversation (Woken & Swales, 1989; Zuengler, 1989a), and the gender of the participants (Leet-Pellegrini, 1980; Varonis & Gass, 1986).

To summarize this review of previous research into the discourse of NS-NNS dyadic interactions, we have seen that this may be characterized by three variables:

(a) The semantic relation between one participant's utterance and the immediately preceding utterance by the other participant (interactional contingency);
(b) The degree to which participants in an interaction are goal-oriented as indicated by the persistence of topics in the discourse; and
(c) The degree to which either participant dominates the discourse by gaining, holding, or ceding the floor, and by controlling the other participant's access to the floor through questions and interruptions.

The organization of discourse created in dyadic interaction is also related to the situational context of the interaction. Different conversational genres such as ritual interactions, interviews, casual conversations, and serious negotiations may produce different structures of discourse. Similarly, the identities and roles of participants, as native or nonnative speakers, as experts or nonexperts, as men or women also affect the structure of the discourse. We turn now to discuss an empirical study of the discourse of one genre of NS-NNS interaction—oral proficiency interviews—and how it relates to the factors we have outlined above.

THE DISCOURSE OF ORAL PROFICIENCY INTERVIEWS

The study that will be reported in the remainder of this paper is motivated by two concerns. The first aim of the study is to provide a description of the discourse of oral proficiency interviews in terms of the three dimensions of discourse structure discussed above—interactional contingency, goal orientation, and dominance. This description will enable us to make a principled comparison between oral proficiency interview discourse and other genres of spoken NS-NNS interaction. In particular, we wish to compare the discourse constructed in oral proficiency interviews with conversational discourse constructed by native and nonnative speakers in non-interview situations. The second aim of the study is to relate variation in the structure of discourse in oral proficiency interviews to features of context. To this end, we will attempt to relate discourse structure to the task in which NS examiner and NNS candidate are engaged, to the theme or subject-matter of the interview, to individual differences among NS examiners, and to the gender of the examiner and candidate. If we understand which contextual factors contribute most to determining the structure of the interview, we may be able to use this information to design more valid instruments for assessing oral proficiency in the future.

Background to the Study

The study is based on oral interviews of the First Certificate in English examination (FCE) of the University of Cambridge Local Examinations Syndicate (UCLES). The FCE is administered in about 70 countries world wide and is taken by approximately 180,000 candidates each year. The present interview format was designed in 1984 and is overtly structured, involving three separate tasks to be performed in a fixed order (UCLES, 1982). Different packs of thematically related materials are made available to examiners, who may choose which pack of materials to use for a given interview.

The interviews generally last 12–15 minutes. The interview format allows several combinations of candidates and examiners to participate in one interview, including two candidates-one examiner, two candidates-two examiners, and three candidates-two examiners. However, the most frequent arrangement is one candidate-one examiner and in this study only the one-to-one combination was considered in order to make for more effective application of the model of dyadic interaction we propose and to facilitate comparison with other studies of dyads.

The data for this study consist of 30 oral interviews, 14 recorded at one testing center in Rio de Janeiro, Brazil, and a further 16 in centers at Rome and Trieste in Italy.

Recording equipment was a small tape recorder controlled by the examiner and in full view of the candidate. All tapes were transcribed by native speakers of British English in Cambridge and checked by the authors. Transcripts were coded by the first author and a native speaking research assistant. All examiners in the study were NSs of British English who had experience interacting with students of English and with the FCE oral interview format. The examiners both conduct the interview and score the candidate's performance on six five-point scales of fluency, grammatical accuracy, sentence pronunciation, the pronunciation of individual sounds, interactive communication, and vocabulary resource. The candidate's final score is the sum of the scores on the six separate scales. There were a total of eight examiners involved, of whom five were males and three were females. All examiners conducted at least two interviews, some conducted more.

Candidates in Brazil were all native speakers of Brazilian Portuguese, while those in Italy were all native speakers of Italian. Candidates' spoken English proficiency as measured by the FCE oral interview ranged from intermediate to advanced. Seven interviews were conducted between female examiners and female candidates, 5 between female examiners and male candidates, 11 between male examiners and female candidates, and 7 between males. In 12 interviews examiners chose the pack of interaction materials focusing on the general theme of "Learning," 7 interviews focused on the theme of "Having a Good Time," and 11 focused on "Work."

Each interview consisted of a warm-up, three separate tasks, and a closing. The warm-up and closing were not included in the analysis and the data consisted solely of the discourse in the three tasks. These were always performed in the same order and consisted of (a) a discussion based on one or more photographs, which we refer to as the *photographs task*; (b) a task in which candidates were asked to relate a printed passage to the photographs they had discussed in the photographs task, which we refer to as the *passages task*; and (c) an activity which involved candidates expressing personal preferences about items in a list of activities related to the theme of the interview. This task we refer to as *ranking activities*.

Discourse Variables

In the present study the basic unit of discourse chosen for analysis is topic. Topic, defined as the person or thing about which something is said, is a suitable discourse construct for analyzing contingency and goal orientation in dyadic interactions since it is possible to identify whether one participant continues or terminates a topic introduced by the other participant. Previous empirical studies of topic continuity in spoken discourse (Givón, 1984) and in written discourse (Schneider & Connor, 1990) have developed the theory of topic continuity and have gone to considerable lengths to describe how, in practice, to identify coreferential topics in succeeding utterances.

It should be noted, however, that there are several limitations in using topic as the basic unit of analysis. First, clearly, participants do not react solely to what their interlocutors say. They may also react in terms of eye contact, posture, and body positioning vis-à-vis their interlocutor as Argyle and Dean (1965) and Erickson and Shultz (1982) have shown. Similarly, as other analysts of face-to-face interaction such as

Gumperz (1979) have revealed, participants may also react not only to what their interlocutors say, but to how they say it—in particular to the intonation of utterances. A full analysis of contingency in dyadic interactions will therefore have to deal with more dimensions of the interaction than are considered here.

Second, as Brown and Yule (1983) and Givón (1984) have indicated, it is often difficult to say whether a given noun phrase or predicate *either* does *or* does not encode the topic of an utterance. Rather, the many constituents of an utterance which are all to a certain extent "given" rather than "new," "topic" rather than "comment" suggest that with topic one is not dealing with an either/or property of discourse but rather a gradable one. Moreover, even if a topic can be identified, what is the scope of the topic? To address these problems, we chose to adapt the topical structure analysis of Schneider and Connor (1990) to spoken data, using the t-unit as the stretch of discourse over which a topic has scope.

Figure 2 about here

The discourse variables measured in this study are defined in Figure 2.[4] Given these measures, one can perhaps distinguish the four interactional styles proposed by Jones and Gerard as follows. Asymmetrical contingency involves more topic initiations by one party than the other and also greater ratification of topics by one party than the other. If the oral interviews exhibit asymmetrical contingency, then examiners will initiate more topics than candidates, while candidates will ratify more examiner-initiated topics than vice-versa. The distinction between reactive and mutual contingencies in the Jones and Gerard model has to do with the degree of goal orientation involved. Reactive contingency involves fairly low topic persistence whereas mutual contingency involves high topic persistence as both parties pursue their internal goals. Reactiveness is high in both cases. Thus, if the interviews exhibit reactive contingency, we are likely to see a large number of topics ratified by the other party, however topic persistence will be low with neither participant predominating. On the other hand if the interviews exhibit mutual contingency, we are likely to see both high degree of mutual ratification of topic initiations as well as high topic persistence.

On the basis of the above discussion and the previous research into NS-NNS discourse summarized above, the following hypotheses were constructed regarding the structure of discourse in the FCE oral interviews.

H1. Since the purpose of the oral interview is to elicit a ratable sample of speech from candidates, candidates will talk more than examiners.

H2. Examiners will dominate the discourse by initiating more topics than candidates.

H3. Oral interviews will exhibit asymmetrical contingency which will be evident because of greater reactiveness by candidates and greater orientation to goals by examiners. Specifically,

(a) Candidates will ratify a greater proportion of topics initiated by their interlocutor than will examiners.

(b) The persistence of topics initiated by examiners will be greater than the persistence of topics initiated by candidates.

The Influence of Context

Our second research question concerns the relationship between discourse in oral proficiency interviews and the social context of those interviews, in particular, who the candidates are speaking to, what they are speaking about, and what task they are engaged in. Our hypotheses are derived from previous findings on variation in phonology, morphology, and syntax in interlanguage speech, which we believe may be equally apparent on the discourse level. We attempt to account for the observed variation in dominance, contingency, and goal-orientation by relating it to the independent contextual factors of interlocutor, theme, task, and gender of participants.

Several studies of interlanguage phonology designed within the general framework of speech accommodation theory have found that the learner's interlocutor plays a significant role in conditioning variation in the learner's second language pronunciation. Beebe (1977) and Beebe and Zuengler (1983) attributed this interlocutor effect to ethnic identification between participants. Berkowitz (1989) found a similar effect and attributed it to the degree of cultural empathy between NS and NNS. Selinker and Douglas (1985) claim to have observed an interlocutor effect in interlanguage discourse in NS-NNS conversations, but this was inseparable in their study from the theme of the conversation. Other studies, however, have cast some doubt on whether accommodation to the speech patterns of an interlocutor is as widespread a phenomenon in other areas of IL as speech accommodation theorists have claimed (Young, 1986, 1991; Zuengler, 1991).

The theme of NS-NNS conversations has been shown to affect learners' pronunciation by Eisenstein and Starbuck (1989), who attribute the effect to learners' emotional investment in the theme of the conversation. Similarly, learners' expertise on a given theme has been shown to affect IL discourse by Selinker and Douglas (1985) and by Zuengler (1989a). Again, however, results in this area are not unequivocal. Smith (1989), for example, found that learners' comprehensibility was *not* different when they were speaking on general themes or when they were speaking on themes within their academic major. In the present study, the theme of the FCE oral interviews was identified as one of the three thematic areas of "Learning," "Having a Good Time," and "Work" according to the pack of materials chosen by the examiner.

Many previous studies of IL variation have established that task plays an important role in conditioning variation in IL phonology (Beebe, 1980; Dickerson & Dickerson, 1977; Gatbonton, 1978; Sato, 1985; Schmidt, 1977; Stølen, 1987; Weinberger, 1987; Wenk, 1986). A smaller number of studies have established a role for task in accounting for variation in IL morphology (Ellis, 1987; Larsen-Freeman, 1975; Tarone, 1985; Tarone & Parrish, 1988) and in the syntax of IL negation (Adamson & Kovac, 1981).

Finally, one aspect of the interlocutor effect which may exert an influence on IL discourse is gender. The combinations of males and females in the interview dyads were examined to see if there was evidence to support previous findings on the differences in

theme-related dominance in NS-NNS interactions between males and between females. Four questions regarding these contextual influences on the structure of the discourse were formulated as hypotheses H4–7.

H4. Dominance, contingency and goal orientation in the discourse of FCE oral interviews (as measured by quantity of talk, number of topic initiations, reactiveness, and topic persistence) will vary according to the examiner.

H5. Dominance, contingency, and goal orientation in the discourse of oral interviews (as measured by quantity of talk, number of topic initiations, reactiveness, and topic persistence) will vary according to the theme of the interview.

H6. Dominance, contingency, and goal orientation in the discourse of oral interviews (as measured by quantity of talk, number of topic initiations, reactiveness, and topic persistence) will vary according to the task in which examiner and candidate are engaged.

H7. Dominance, contingency, and goal orientation in the discourse of oral interviews (as measured by quantity of talk, number of topic initiations, reactiveness, and topic persistence) will vary according to the gender of examiner and candidate.

The analytical procedures for testing the seven hypotheses H1–7 and the results are discussed in the following section.

ANALYSIS AND RESULTS

Simple univariate statistical tests on the distribution of the discourse variables convinced us that the data are not normally distributed. Standard deviations are high in all cases and the range between maximum and minimum values is also large. More systematic tests revealed that the distributions of most variables are positively skewed and sharply peaked. In addition, the data do not represent a random sample of the population. On the basis of these indications, we consider means and standard deviations to be inappropriate in reporting the results. Instead, medians are more appropriate measures of central tendency and nonparametric procedures based on ranks rather than raw scores are used to investigate relations among variables, as recommended by Hatch and Lazaraton (1991).

Table 1 about here

Table 1 shows the results of four Wilcoxon Rank Sums tests performed to test hypotheses H1–3 regarding differences between examiner and candidate discourse. As Table 1 shows, the median quantity of talk by candidates in any one task was 21 t-units, almost twice the median amount of talk by examiners (12.5 t-units). This difference is as predicted by hypothesis H1 and is highly significant ($p=.0001$). However, the median number of topics initiated by the examiner and by the candidate did not differ significantly. Thus hypothesis H2, which predicted that examiners will control the discourse through more topic initiations, is not supported.

Dominance by the examiner *is* noticeable, however, in the degree of reactiveness of candidate and examiner to each other's topic initiations. Examiners ratified a median of only 33% of topics initiated by the candidate, while candidates' ratified twice as many (67%) topics initiated by the examiner. This is the difference predicted by hypothesis

H3(a) and is highly significant (p=.0001). With regard to the degree to which either party is oriented to internal goals, which we have measured by the persistence of topics they initiate, results show a small but nonetheless significant difference (p=.0116). The median persistence of topics initiated by candidates was 4 t-units, while the median persistence of topics initiated by examiners was 4.5 t-units. This is the difference predicted by hypothesis H3(b). Taken together, these two results show that FCE oral interview discourse is indeed characterized by asymmetrical contingency. However the asymmetry lies primarily in the degree to which the examiner and candidate react to topics initiated by each other and, to a lesser extent, in the degree to which the examiner is more goal-oriented than the candidate.

Table 2 about here

Table 2 shows the median scores for the eight discourse variables related to the examiner in the interview together with results of Kruskal-Wallis tests performed on each variable in order to investigate the relations between examiner and discourse predicted by hypothesis H4. Results show that the examiner does not significantly influence the amount of talk, the number of topics initiated by either party, nor the persistence of topics in the discourse. Examiners do not vary in their reactivity to candidates' topic initiations; however, as the *RET* column of Table 2 shows, candidates *do* vary significantly in the proportion of topics they ratify from different examiners (p=.0028).

Further analysis of the candidate's reactiveness to examiner-initiated topics (*RET*) was carried out by means of Ryan's procedure (Linton & Gallo, 1975) in order to identify which examiners caused greater reactiveness among candidates.[5] Ryan's procedure locates the significant differences among specific pairs of medians when a Kruskal-Wallis test has shown a variable with three or more levels to contribute significantly to variation in the dependent variable. In the case of *RET* related to examiner, Ryan's procedure did not find any significant differences among the reactiveness of candidates with different individual examiners.

In order to evaluate the strength of the association between examiner and topic ratifications by the candidate, an η^2 test of association was performed (Hatch & Lazaraton, 1991). This test showed that only 24% of the variation in candidates' reactiveness can be accounted for by the examiner. There are thus clearly other, as yet unidentified, factors at work in influencing the reactiveness of candidates to examiners' topic initiations and we would be wise not to overestimate the role played by the examiner.

Table 3 about here

Table 3 shows results of the investigation of the relationship between interview discourse and the theme of the interview predicted by hypothesis H6. The investigation of hypotheses H5-7 was carried out in a similar way to that reported for hypothesis H4. Tables 3, 4, and 5 report median values of the discourse variables and the results of Kruskal-Wallis tests in a similar way to Table 2.

Table 3 shows that the theme of the interview, as determined by the choice of the interaction pack, had a significant effect on only one discourse variable—the persistence of topics initiated by the examiner. Analysis of the difference among medians by Ryan's procedure shows that the theme of "Work" produced significantly longer topics than the theme of "Learning" ($p<.05$) but that the length of examiner-initiated topics in interviews on the theme of "Having a Good Time" did not differ significantly from interviews on either of the other two themes. The median persistence of examiner-initiated topics in interviews on the theme of "Work" was 6 t-units, while the median persistence of examiner-initiated topics in interviews on the theme of "Having a Good Time" was 4.5 t-units and "Learning" was 4 t-units. However, an η^2 value of 10% shows that there is a weak association between theme and topic persistence and that other factors must also be influencing the variation.

Table 4 about here

The influence of task on the discourse is shown in Table 4. Task appears to have the strongest effect of any of the contextual variables considered and its effect is most apparent in the measures of dominance. The quantity of talk from both examiner and candidate, the number of topics initiated by both parties, and the persistence of topics initiated by the candidate are all affected by the task in which examiner and candidate are engaged, as predicted by hypothesis H6. However, no significant relationship exists between task and the two measures of contingency, nor the persistence of topics initiated by the examiner. The results of Ryan's procedure to locate the differences among medians show that on the measures of quantity and topic initiations, all three tasks differed significantly from each other ($p<.05$), with the ranking activities producing the greatest quantity of talk and the greatest number of topic initiations; the photographs task producing the least; and the passages task producing an amount intermediate between the other two. As far as the persistence of topics initiated by the candidate is concerned, the photographs task produced significantly longer topics (a median of 5 t-units) than the other two. There was no significant difference in the persistence of candidate-initiated topics produced by the ranking activities or the passages task.

Tests of association by means of the η^2 statistic between task and the five discourse variables on which it exerted a significant effect show a moderate to strong association with quantity of examiner and candidate talk (50% and 57% respectively), a weaker association with the number of examiner and candidate topic initiations (35% and 42% respectively), and a weak association with the persistence of candidate-initiated topics (21%). It should be recalled at this point that in the format of the interview, the

three tasks always occurred in the same order within one interview, namely, first photographs, second passages, and last ranking activities. It is thus impossible to separate the effect of task from a serial effect. It may well be that by the third task, the candidate and examiner are more relaxed than they were in the early stages of the interview and are thus able to speak more freely and initiate more topics than in the previous tasks. Without further studies in which the tasks are presented in random order, it is not possible to say whether the manifestly different discourse produced on the three tasks is due to the nature of the tasks themselves or to a serial effect or, as is most likely, to a combination of the two.

Table 5 about here

Finally, we turn to the effect of the gender of the participants on the interview discourse predicted by hypothesis H7. Table 5 shows that the gender of examiner and candidate significantly affect two aspects of the discourse—the persistence of examiner-initiated topics ($p=.0036$) and their ratification by the candidate ($p=.0008$). Gender did not significantly affect quantity or number of topic initiations from either party, neither did it affect the persistence of candidate-initiated topics nor their ratification by the examiner.

Investigating now the differences among medians, Ryan's procedure shows that interviews between women produce significantly more reactiveness by candidates (82%) than interviews between male examiners and female candidates (64%) ($p<.05$). Interviews between female examiners and male candidates and those between two men did not differ significantly from the other gender combinations on this variable. As regards the persistence of examiner-initiated topics, again the interviews between women produced significantly longer topics (a median of 6.2 t-units) than interviews between a male examiner and a female candidate (4 t-units) or between two men (3.7 t-units) ($p<.05$). The persistence of topics initiated by a female examiner in interviews with a male candidate did not differ significantly from that produced by the other gender pairings. However, η^2 tests of association reveal that only 15% of the variance in the reactiveness of candidates and 19% of the variance in the persistence of examiner-initiated topics is attributable to gender. Other factors in addition to gender are clearly at work here. The significance of these results will be discussed in the remaining sections.

DISCUSSION

These results highlight several interesting features of NS-NNS interactions in oral proficiency interviews. First of all, the analysis confirms what scholars had predicted about interview discourse. The NS examiner and NNS candidate make radically different contributions to the discourse. The degree to which the examiner exercises control over the discourse is apparent from the much greater reactiveness of candidates to examiners' topic initiations than vice versa, and from the slightly longer persistence of examiner-initiated topics. The discourse we observe in these interviews shows evidence of greater goal orientation by one party and much greater reactiveness by the other party. It thus clearly exemplifies the style that Jones and Gerard have called asymmetrical contingency.

However, when we look at the measures of dominance—quantity of talk and number of topic initiations—the data appear to suggest that the candidate dominates the interview since the candidate talks about twice as much as the examiner, while there is no significant difference between the number of topics initiated by either party. These measures of dominance do not capture the underlying control over the right to speak since apparently the candidate has the right to speak more than the examiner. However, the results for reactiveness and persistence show that, nonetheless, this right is limited by the examiner. According to the rules of the interview, the examiner gives the floor to the candidate most of the time but still controls what is spoken about and is able to take or withhold the floor at any time. The greater reactiveness of the candidate is evidence that this control over the discourse is accepted. And thus, perhaps, in the context of these interviews it is more appropriate to talk of a candidate's obligation (rather than right) to speak.

Jones and Gerard and van Lier claim that conversation involves more balanced reactiveness by both parties to each other's topic initiations, and McLaughlin (1984) defines conversation as "relatively informal social interaction in which the roles of speaker and hearer are exchanged in a nonautomatic fashion under the collaborative management of both parties" (p. 271). However, since this is the first study to operationalize this model, we still lack an analysis of baseline NS-NNS conversational data which would allow us to make a true comparison. A particular strength of the contingency model has been in demonstrating that dominance in these interviews is not simply reflected by the relative amount of talk by participants nor by the relative numbers of topic initiations, as several earlier studies of dominance had claimed. Instead, dominance can be seen in these interviews in the degree of reactiveness of the candidate to the examiner's initiations, but not vice versa.

If we turn our attention now to the contextual factors which affect the structure of the discourse, we see that different contextual factors affect participants differently and this may be accounted for by participants' different roles in the interview. We will first consider which contextual factors affect only candidate discourse; we will then go on to consider which contextual factors affect only examiner discourse; finally, we will consider which contextual factors affect both candidate and examiner equally.

Contextual Influences on the Candidate

The reactiveness of the NNS candidate to topics initiated by the NS examiner appears from Tables 2 and 5 to be related to certain individual differences among examiners. Female candidates are more reactive to topics initiated by female examiners and all candidates appear to vary in the degree to which they ratify topics initiated by different examiners. There is probably some interaction between candiadtes' reactiveness due to examiner and that due to gender here since examiners H, E, and F, who induced the most reactiveness from candidates, are all female, while examiners A, B, G, C, and D are male and they induced the least reactiveness. However, other individual differences between examiners may be relevant here such as age, empathy, and other characteristics which lead to greater solidarity between examiner and candidate. These characteristics were not measured in this study

The persistence of topics initiated by the NNS candidate appears to be due to task. The photographs task and the ranking activities produce candidate-initiated topics which persist about twice as long as topics initiated by the candidate in the passages task. What makes the passages task so different from the other two? This task involves the candidate matching a written text with a photograph and is unique in certain respects since it does not allow for a negotiated outcome (Rulon & McCreary, 1986; Young, 1984, 1988). In the passages task there is only one right answer and this is known in advance to the examiner—the correct match of text and photograph. The examiners' topic initiations in the passages task are thus display questions while in the other two tasks in the interview the examiners pose very few display questions.[6] In the other two tasks there are no predetermined outcomes and candidates are thus able to negotiate the outcome with examiners. Previous research on ESL classroom discourse by Brock (1986) has shown that students' responses to teachers' display questions are significantly shorter than responses to referential questions and a similar effect is clearly visible in the present results.

Contextual Influences on the Examiner

The examiner's contribution to the discourse appears to be affected by contextual factors in ways different from the candidate. The examiner's degree of reactiveness is not affected by any contextual factors. However, the goal orientation of the examiner appears to be affected by gender as well as by the theme of the interview. The effect of gender on examiner topic persistence may be related to the reactiveness of the candidate, since both the persistence of examiner-initiated topics and candidates' reactiveness to them are significantly higher only in interviews between women. Since topic persistence is counted across turns by both participants, it will be higher if one participant ratifies the topics introduced by the other as appears to be the case here. On the other hand, if an examiner-initiated topic is not ratified by the candidate the topic dies and topic persistence will have a value approaching unity. Thus greater reactiveness by the candidate produces longer topics initiated by the examiner.

The theme of "Work" produces significantly longer examiner topics than the theme of "Learning." This thematically related difference in the discourse may be due to the fact that the candidates in the interviews were mostly young adults, many of whom

were looking forward to graduating soon from college or university. A favorite set of topics for discussion in many interviews with these candidates centered on their future plans for a career after graduating. Examiners often initiated topics in this area by asking candidates to talk about their career plans. Such topics persisted in the discourse because candidates had a lot of ideas and opinions about the subject. The emotional investment of the candidates in the theme of work at this particular stage in their lives may explain the link between theme and topic persistence. As has been mentioned above, previous studies have shown a relationship between a NNS's expertise on a given theme or emotional investment in the theme and the structure of NS-NNS interactions. The present study did not measure independently candidates' expertise or emotional investment in the three themes and we are thus only able to speculate on the reasons for the apparent effect of theme on the discourse. We look forward to further studies of oral proficiency interviews which investigate this relationship in greater detail.

Contextual Influences on the Discourse as a Whole

Of all the contextual factors that were investigated, task exerts the strongest influence on the structure of the discourse in these interviews. A great deal of the variance in the quantity of talk from both parties is due to task. Variance in the number of topic initiations by both parties is also influenced to a significant degree by task, albeit less so than with quantity. Moreover, the influence of each of the three tasks is significantly different one from the other, with the ranking activities producing the greatest quantity of talk and number of topic initiations, the passages task producing the least, and the photographs task producing an amount intermediate between the other two. We have already discussed the possible reasons why the passages task stands out as producing the shortest topics and the results for quantity and topic initiations may be explained in the same way. The passages task does not allow candidates to negotiate the outcome of the task and results in display questions from the examiner. Both these factors combine to create very short interactions.

However, what causes the greater quantity and higher number of topic initiations on the ranking activities in comparison to the photographs task? We believe these results are due both to a serial effect and to the nature of the ranking activities task. The ranking activities are the last task in the interview and the photographs task is the first task. As we have suggested above, by the later stages of the interview examiner and candidate may be more relaxed in each other's company and this may lead both parties to speak more freely and at greater length. In addition, the ranking activity is the only one in which candidates are asked to talk about their own beliefs and their own experiences. While this may also happen to a certain degree in discussing the photographs, that task and the passages task center primarily on the discussion of items which are elicited by the interaction pack materials which candidates have never seen before the interview. In the third activity, therefore, candidates may simply have more to talk about than in the other two and this is reflected in the higher values for quantity and topic initiations on this task.

Finally, it should be noted that task, while exerting a strong influence on the quantity of talk and the number of topic initiations by both parties as well as on the

persistence of candidate-initiated topics, does not exert a significant effect on the reactiveness of either the candidate or the examiner. This suggests an interesting distinction between, on the one hand, personal contextual variables such as gender and other individual differences among examiners and, on the other hand, situationally-related contextual variables such as task and theme. Discourse in these interviews is mediated by permanent personal characteristics of examiners and candidates such as gender, and by temporary situational factors such as task and theme. The reactiveness of a candidate in ratifying topics initiated by the examiner depends on who the candidate is speaking to, while the quantity of candidate talk, the number of topics initiated by the candidate and their persistence depends on situational factors not related to the person to whom she or he is speaking. In other words, reactiveness appears to be interpersonal, while the other features of the discourse appear to be situational in origin.[7]

CONCLUSIONS

This study is principally exploratory and descriptive. There are, to our knowledge, remarkably few descriptive linguistic studies of discourse in oral proficiency interviews.[8] Given our present limited knowledge of oral interview discourse, this study is also limited in a number of ways. First, it describes only one oral proficiency interview format—the Cambridge FCE—and conclusions from this study may not be applicable to other interview formats. Second, the study is descriptive rather than experimental since we simply took a non-random sample of existing interviews to study and we have no knowledge of several of the potentially relevant variables in the study, such as learners' expertise or emotional investment in the theme of the interview, examiners' empathy or age, or the potential confound between task and a serial effect. Results from the study should therefore be treated as suggestive rather than conclusive. Finally, although the contingency model of dyadic interaction has very effectively summarized the overall structure of the discourse, it is nonetheless a static rather than a dynamic description. That is to say, it does not reveal the process by which discourse is *created* through interaction but only gives us a snapshot of the product of that process. We need to go into much greater detail in describing the nature of contingency as it exists between each set of turns in order to understand how discourse is created.

We believe the strengths of the study lie in describing the discourse of oral proficiency interviews in terms of a general model of dyadic interaction because in so doing we have illuminated some important properties of the discourse. Dominance in these dyadic interactions is not simply a question of speaking more or of initiating more topics but has to do rather with the goal orientation of the examiner and the reactiveness of the candidate. Dominance is also evident in the asymmetry of the interaction: the examiner is far less reactive than the candidate and the candidate appears a little less goal-oriented than the examiner. We may assume, following Jones and Gerard and van Lier, that this kind of discourse is a unique product of the interview situation. Conversations, as defined by McLaughlin (1984) and van Lier (1989) involving NNSs in non-testing situations, may be very different in structure from the discourse of oral proficiency interviews. If this is the case, then van Lier's (1989) witty characterization of the peculiarities of oral interview discourse as "reeling, writhing, drawling, stretching, and

fainting in coils" (p. 489) is apt indeed. The oral proficiency interviews that we have examined here bear very little resemblance to the collaborative management of talk by both parties that we believe to be the structure of non-testing conversations. The interviews often fail to elicit discourse from NNS candidates such as maintaining a topic over a large number of turns that we may want to consider in assessing oral proficiency. At the same time, the interviews often confront candidates with situations such as having to respond to rapidly shifting topics initiated by the NS, which under other circumstances we may consider of little relevance in the assessment of oral proficiency.

We have also seen that there exists significant variation in the discourse. We attribute this variation to the effect of the examiner, the task, the theme of the interview, and the gender of the participants. Since task appears to contribute most to the structure of the discourse, it is probably a good idea that the FCE interviews sample candidates' performance over a range of three different tasks. Still, task is not all. There are some interpersonal features of the candidate's discourse such as reactiveness which appear to be unaffected by task but sensitive instead to the learner's interlocutor. Gender appears to play a role in this, but it would be worthwhile to conduct further investigations into examiners' interactional styles to see what other factors cause greater or lesser reactiveness among candidates.

Yet, for all this detailed description of interview discourse and the variation within it, we still do not know whether any of this makes a difference to the examiner's assessment of the learner's oral proficiency. Now that we have a working model of interview discourse which both describes and explains important aspects of the linguistic interaction, we are in a good position to see whether variations in the discourse affect examiners' judgments. We anticipate that further research along these lines will help us to understand what native speakers react to when they judge the oral proficiency of a learner of their language and, most interestingly, the nature of oral proficiency itself.

NOTES

1.For the sake of terminological clarity we wish to distinguish between two senses of the term "topic." Unfortunately, conventional definitions of topic as "what is talked about" (Richards, Platt, & Weber, 1985) do not specify the stretch of utterance to which a topic applies and this has led to considerable terminological confusion over what is already a difficult term to define and use. In this paper we will continue to use the term "topic" to refer to what is talked about in one t-unit or clause. We will, however, use the term "theme" to refer to what is talked about in a longer stretch of discourse such as one of the interviews. Thus we will refer to "this picture" as the topic of the utterance "Tell me about this picture," while we will refer to "Learning," "Having a Good Time," or "Work" as the theme of the interviews discussed below. Topic, as used in this paper, thus corresponds loosely to the sentential topic described by Brown and Yule (1983), while theme corresponds loosely to what Brown and Yule call the title of a stretch of discourse.

2. We use the term "oral proficiency interview" in this paper in its generic sense to mean an interview format designed to measure spoken ability in a second language. We are aware, however, that in some circles the term has acquired the more restrictive meaning of an interview test which uses the functionally-based language rating scales designed by the Interagency Language Roundtable (Lowe, 1982) and widely disseminated in the *ACTFL Proficiency Guidelines* (American Council on the Teaching of Foreign Languages, 1986). The oral proficiency interview described in this paper differs in several important respects from the ILR/ACTFL format.

3.Jones and Gerard use the term "planning" to refer to the degree to which each participant in an interaction attempts to achieve his or her own goals. Planning is a term which has become widely used in second language acquisition in the sense in which it was used by Ochs (1979) to refer to forethought and organizational preparation in advance of speech. While planning in Ochs' sense surely coincides with the goal orientation of the Jones and Gerard model, goal-orientation in interaction does not necessarily imply forethought and advance preparation. Thus, to avoid possible confusion, "goal-orientation" will be used in this paper to refer to what Jones and Gerard have called "planning."

4.Our research questions are (a) which participant in the interviews—examiner or candidate—is more dominant, reactive, or goal oriented; and (b) how much of the variance in dominance, reactiveness, and goal orientation can be attributed to the contextual factors of interlocutor, theme, task, and gender. For the purposes of statistical analyses which are congruent with these research questions, dominance, contingency, and goal orientation are measured on interval scales. The units of the scales are t-units in the case of quantity and persistence of topics, and topics in the case of topic initiation. Ratification of topics is measured on a ratio scale. That is to say, our variables are *not* t-units or topics, the measurement of which would be more appropriately done on a scale of frequency.

5.Ryan's procedure is a nonparametric test for locating differences among medians. It is analogous to the Scheffé multiple-range test for means found to be significant by parametric analysis of variance. See Hatch and Lazaraton (1991, pp. 337-340).

6.Question prompts are used by examiners in all stages of the interviews and are much more frequent in all tasks than prompts in which the examiner simply requests the candidate to talk about X. Therefore, the faster decay of topics in the passages task does not appear to be due to the examiner using a question prompt in this task rather than a "talk about X" prompt. Rather, the discourse function of the question prompts in the passages task differs from that of question prompts during the other two tasks. In the passages task, the question seeks to elicit information that is already known to the examiner, while in the other two tasks the examiner's questions are more often genuinely information-seeking.

7.The distinction between interpersonal and situational causes of discourse variation in these interviews suggests ways in which the reliability of the interviews as measures of oral proficiency may be improved by eliciting similar discourse across different examiners and tasks. Training of examiners specifically directed toward encouraging candidate reactiveness is likely to be an effective way of increasing reliability. Equally, careful design of the tasks to be performed in the interview and choice of themes for the thematically-based interaction packs may result in greater standardization of both quantity of talk and topic initiations by both parties.

8.Recent work by researchers such as Lazaraton (1991) and Yoshida (1991) working within theoretical frameworks different from the one presented here promises to greatly increase our understanding of the discourse of oral proficiency interviews.

REFERENCES

Adamson, H. D., & Kovac, C. (1981). Variation theory and second language acquisition: An analysis of Schumann's data. In D. Sankoff & H. J. Cedergren (Eds.), *Variation omnibus* (pp. 285-292). Edmonton, Alberta: Linguistic Research.

American Council on the Teaching of Foreign Languages. (1986). *ACTFL proficiency guidelines*. Hastings-on-Hudson, NY: Author.

Argyle, M., & Dean, J. (1965). Eye contact, distance and affiliation. *Sociometry, 28*, 289–304.

Beebe, L. M. (1977). The influence of the listener on code-switching. *Language Learning, 27*, 331-339.

Beebe, L. M. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*, 433-448.

Beebe, L. M., & Zuengler, J. (1983). Accommodation theory: An explanation for style shifting in second language dialects. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 195-213). Rowley, MA: Newbury House.

Berkowitz, D. (1989). The effect of cultural empathy on second-language phonological production. In M. R. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language variation* (pp. 101-114). New York: Plenum.

Brock, C. A. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly, 20*, 47-59.

Brown, G., & Yule, G. (1983). *Discourse analysis*. New York: Cambridge University Press.

Dickerson, L. J., & Dickerson, W. B. (1977). Interlanguage phonology: Current research and future directions. In S. P. Corder & E. Roulet (Eds.), *Actes du 5ème Colloque de Linguistique Appliquée de Neuchâtel: The notions of simplification, interlanguages and pidgins and their relation to second language learning* (pp. 18-29). Geneva: Droz.

Douglas, D., & Selinker, L. (1985). Principles for language tests within the 'discourse domains' theory of interlanguage: Research, test construction and interpretation. *Language Testing, 2*, 205-226.

Eisenstein, M. R., & Starbuck, R. J. (1989). The effect of emotional investment on L2 production. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition volume II: Psycholinguistic issues* (pp. 125-137). Clevedon, Avon: Multilingual Matters.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9*, 1-20.

Erickson, F., & Shultz, J. (1982). *The counselor as gatekeeper: Social interaction in interviews*. New York: Academic Press.

Gaies, S. J. (1982). Native speaker-nonnative speaker interaction among academic peers. *Studies in Second Language Acquisition, 5*, 74–81.

Gass, S. M., & Varonis, E. M. (1986). Sex differences in nonnative speaker-nonnative speaker interactions. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 327-351). Rowley, MA: Newbury House.

Gatbonton, E. (1978). Patterned phonetic variability in second language speech: A gradual diffusion model. *Canadian Modern Language Review, 34*, 335-347.

Givón, T. (1984). Universals of discourse structure and second language acquisition. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 109-136). Philadelphia: John Benjamins.

Grice, H. P. (1989). Logic and conversation. In H. P. Grice, *Studies in the way of words* (pp. 3-143). Cambridge, MA: Harvard University Press.

Gumperz, J. J. (1979). The retrieval of sociocultural knowledge in conversation. *Poetics Today, 1*, 273–286.

Hatch, E. (1983). Simplified input and second language acquisition. In R. W. Andersen (Ed.), *Pidginization and creolization as language acquisition* (pp. 64-86). Rowley, MA: Newbury House.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.

Jones, E. E., & Gerard, H. B. (1967). *Foundations of social psychology*. New York: Wiley.

BEST COPY AVAILABLE

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409-419.

Lazaraton, A. (1991). *A conversation analysis of structure and interaction in the language interview.* Unpublished Ph.D. dissertation, University of California at Los Angeles.

Leet-Pellegrini, H. M. (1980). Conversational dominance as a function of gender and expertise. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives* (pp. 97-104). New York: Pergamon.

Linton, M., & Gallo, P. S. (1975). *The practical statistician: A simplified handbook of statistics.* Monterey, CA: Brooks/Cole.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*, 126-141.

Lowe, P., Jr. (1982). *ILR handbook on oral interview testing.* Washington, DC: Defense Language Institute.

McLaughlin, M. L. (1984). *Conversation: How talk is organized.* Beverly Hills, CA: Sage.

Ochs, E. (1979). Planned and unplanned discourse. In T. Givón (Ed.), *Syntax and semantics. Volume 12: Discourse and semantics* (pp. 51-80). New York: Academic Press.

Richards, J., Platt, J., & Weber, H. (1985). *Longman dictionary of applied linguistics.* Harlow, Essex: Longman.

Rulon, K. A., & McCreary, J. (1986). Negotiation of content: Teacher-fronted and small-group interaction. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 182-199). Rowley, MA: Newbury House.

Sato, C. J. (1985). Task variation in interlanguage phonology. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 181-196). Rowley, MA: Newbury House.

Schegloff, E. A. (1978). On some questions and ambiguities in conversation. In W. U. Dressler (Ed.), *Current trends in textlinguistics* (pp. 81–102). New York: Walter de Gruyter.

Schmidt, R. W. (1977). Sociolinguistic variation and language transfer in phonology. *Working Papers in Bilingualism, 12*, 79-95.

Schneider, M., & Connor, U. (1990). Analyzing topical structure: Not all topics are equal. *Studies in Second Language Acquisition, 12*, 411-427.

Selinker, L., & Douglas, D. (1985). Wrestling with 'context' in interlanguage theory. *Applied Linguistics, 6*, 190-204.

Selinker, L., & Douglas, D. (1989). Research methodology in contextually-based second language research. *Second Language Research, 5*, 93-126.

Smith, J. (1989). Topic and variation in ITA oral proficiency: SPEAK and field-specific tests. In R. Young (Ed.), The training of international teaching assistants [Special issue]. *English for Specific Purposes, 8*, 155-167.

Stølen, M. (1987). The effect of affect on interlanguage phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language system* (pp. 389-400). Cambridge, MA: Newbury House.

Takahashi, T. (1989). The influence of the listener on L2 speech. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition. Volume I: Discourse and pragmatics* (pp. 245-279). Clevedon, Avon: Multilingual Matters.

Tarone, E. E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning, 35*, 373-404.

Tarone, E. E. (1988). *Variation in interlanguage.* London: Edward Arnold.

Tarone, E. E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning, 38*, 21-44.

University of Cambridge Local Examinations Syndicate. (1982). *Cambridge examinations in English: Changes of syllabus in 1984.* Cambridge: Author.

van Lier, L. (1989). Reeling, writhing, drawling, stretching and fainting in coils: Oral proficiency interviews as conversation. *TESOL Quarterly, 23*, 489-508.

Weinberger, S. H. (1987). The influence of context on syllable simplification. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language system* (pp. 401-417). Cambridge, MA: Newbury House.

Wenk, B. (1986). Crosslinguistic influence in second language phonology: Speech rhythms. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 120-133). New York: Pergamon.

Woken, M. D., & Swales, J. (1989). Expertise and authority in native—non-native conversations: The need for a variable account. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition. Volume I: Discourse and pragmatics* (pp. 211-227). Clevedon, Avon: Multilingual Matters.

Yoshida, Y. (1991). *The use of communication strategies use by adult Japanese speakers of English as a second language in language proficiency interviews.* Unpublished master's thesis, Southern Illinois University at Carbondale.

Young, R. (1984). Negotiation of outcome and negotiation of meaning in ESL classroom interaction. *TESOL Quarterly, 18,* 525-526.

Young, R. (1986). The acquisition of a verbal repertoire in a second language. *Penn Working Papers in Educational Linguistics, 2*(1), 85-119.

Young, R. (1988). Computer-assisted language learning conversations: Negotiating an outcome. *CALICO Journal, 5*(3), 65-83.

Young, R. (1991). *Variation in interlanguage morphology.* New York: Peter Lang.

Zuengler, J. (1989a). Assessing an interaction-based paradigm: How accommodative should we be? In M. R. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language variation* (pp. 49-68). New York: Plenum.

Zuengler, J. (1989b). Performance variation in NS-NNS interactions: Ethnolinguistic difference, or discourse domain? In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition. Volume I: Discourse and pragmatics* (pp. 228-244). Clevedon, Avon: Multilingual Matters.

Zuengler, J. (1991). *Accommodation theory and the contrastive analysis hypothesis: Both better as weak versions than strong.* Paper presented at the Second Language Research Forum, February 28-March 3, University of Southern California, Los Angeles.

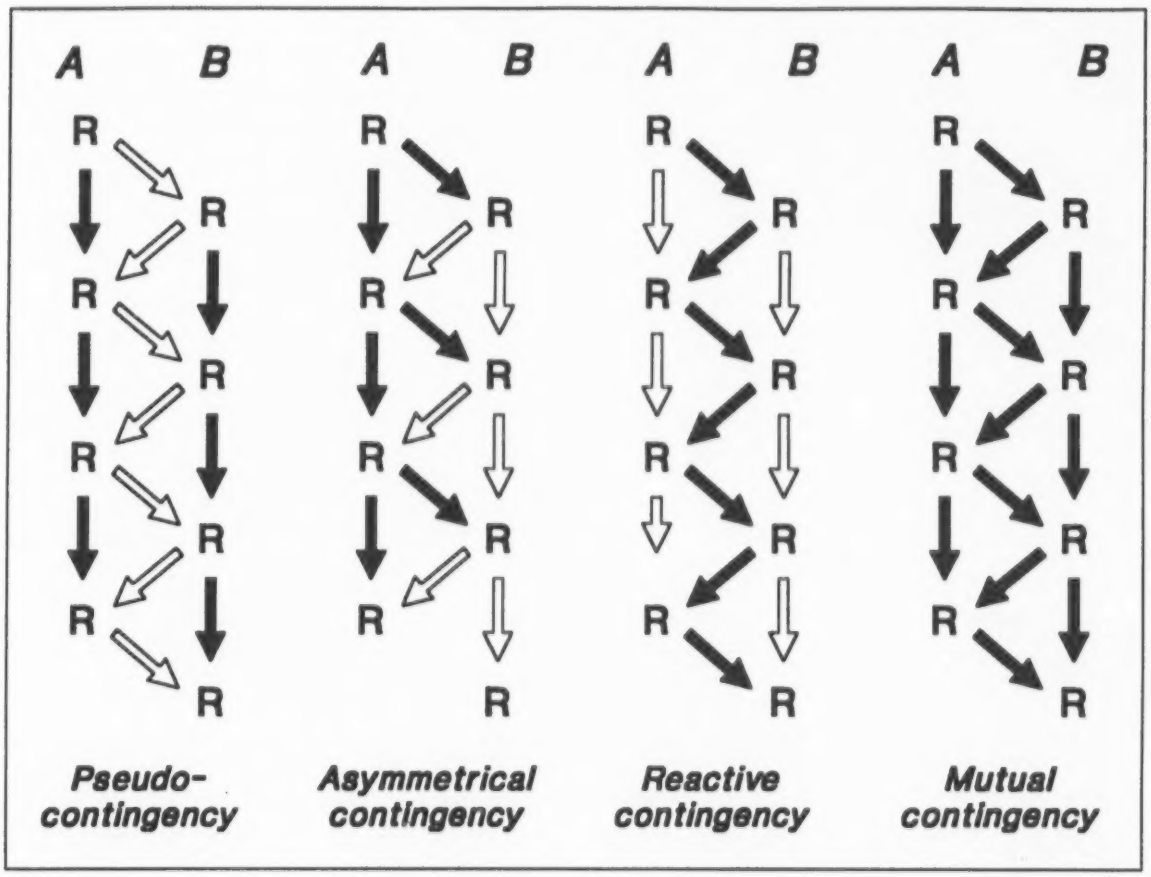

Figure 1. Styles of dyadic discourse related to contingency and orientation of speakers to internal goals. Based on Jones & Gerard (1967, p. 507).

Measures of dominance

Examiner Quantity (EQ): Quantity (jn t-units) of the examiner's talk during one task.
Candidate Quantity (CQ): Quantity (in t-units) of the candidate's talk during one task.

Measures of dominance and contingency

Examiner Topic Initiations (ET): Number of topics initiated by the examiner during one task.
Candidate Topic Initiations (CT): Number of topics initiated by the candidate during one task.

Measures of contingency

Ratification by Candidate of Examiner-initiated Topics (RET): The proportion of topics initiated by the examiner which are the topics of subsequent turns by the candidate (expressed as a percentage).

Ratification by Examiner of Candidate-initiated Topics (RCT): The proportion of topics initiated by the candidate which are the topics of subsequent turns by the examiner (expressed as a percentage).

Measures of goal orientation

Persistence of Examiner-initiated Topics (PET): Mean persistence (in t-units) of all topics initiated by the examiner during one task. The persistence of a given topic is measured by counting total t-units across both speakers from first to last mention of the topic.

Persistence of Candidate-initiated Topics (PCT): Mean persistence (in t-units) of all topics initiated by the candidate during one task.

Figure 2. Discourse variables.

Table 1. Differences between examiner and candidate discourse in 30 FCE oral interviews

	Quantity of Talk by		Topic Initiations by		Percentage of Topics Initiated by Which Are Ratified		Persistence of Topics Initiated by	
	Ex.	Cand.	Ex.	Cand.	Ex.	Cand.	Ex.	Cand.
	EQ	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
N	90	90	90	90	90	87	90	87
Median	12.5	21	5	7	67%	33%	4.5	4
Mean Rank	72.7	108.3	85.7	95.3	120.1	56	98.6	79.1
Wilcoxon Z	–4.57		–1.25		–8.23		–2.52	
Prob > |Z|	.0001		.2133		.0001		.0116	
Ex.-Cand. differences	*CQ > EQ*		Not significant		*RET > RCT*		*PET > PCT*	

N is the number of tasks. Each interview contained three tasks, for a total of 90 observations. However, in the case of three tasks the candidate did not initiate any topics, making the values of *RCT* and *PCT* meaningless. Thus, only 87 observations are reported for these two variables. Significance level set at $\alpha=.05$; df=1.

Table 2. Median scores and Kruskal-Wallis tests for interview discourse related to examiner

Examiner	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
A	12	15.5	21	7	5.5	75%	50%	4.8	4.6
B	21	9	21	4	7	71%	22%	3.5	4
C	9	13	21	4	5	50%	30%	4	3.8
D	6	16.5	17	5	8	45%	20%	3.9	3.8
E	12	18	26.5	8	8	77%	36%	5.7	3.5
F	15	12	16	4	5	75%	33%	6	4
G	6	14	23.5	6.5	8	59%	40%	3.8	2.3
H	9	10	19	6	4	83%	25%	5.3	4.7
Kruskal-Wallis χ^2		9.90	4.25	7.67	4.46	21.72	13.25	13.40	6.29
p		n.s.	n.s.	n.s.	n.s.	.0028	n.s.	n.s.	n.s.
η^2 test of association						.24			

Significance level set at α=.05; df=7. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums). Median scores are shown in Tables 2–5 rather than rank sums because they are more informative.

Table 3. Median scores and Kruskal-Wallis tests for interview discourse related to theme

Theme	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
Learning	36	14.5	21	5	6.5	66%	33%	4	4
Having a Good Time	21	11	18	6	8	75%	25%	4.5	3.6
Work	33	12	21	5	6	71%	33%	6	4.4
Kruskal-Wallis χ^2		.36	.65	1.14	3.69	2.04	.53	9.25	.50
p		n.s	n.s	n.s.	n.s.	n.s.	n.s.	.0098	n.s.
η^2 test of association								.10	

Significance level set at α=.05; df=2. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).

Table 4. Median scores and Kruskal-Wallis tests for interview discourse related to task

Task	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
Photographs	30	12	21	5	7	77%	40%	4.7	5
Passages Task	30	7	10	3	2.5	67%	20%	4.1	2.5
Ranking Activities	30	20.5	35.5	8	10.5	67%	32%	5	4
Kruskal-Wallis χ^2		44.12	50.68	31.48	37.70	4.34	5.62	.73	18.51
p		.0001	.0001	.0001	.0001	n.s.	n.s.	n.s	.0001
η^2 test of association		.50	.57	.35	.42				.21

Significance level set at α=.05; df=2. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).

Table 5. Median scores and Kruskal-Wallis tests for interview discourse related to gender

Gender Ex	Cand	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
F	F	21	16	19	6	5	82%	33%	6.2	5.5
	M	15	12	21	5	7	67%	29%	4.8	3
M	F	33	13	22	5	8	64%	33%	4	4
	M	21	11	19	4	6	67%	24%	3.5	3.7
Kruskal-Wallis χ^2			3.40	2.29	2.76	1.99	13.57	4.23	16.79	3.40
p			n.s.	n.s.	n.s.	n.s.	.0036	n.s.	.0008	n.s.
η^2 test of association							.15		.19	

Significance level set at α=.05; df=3. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).

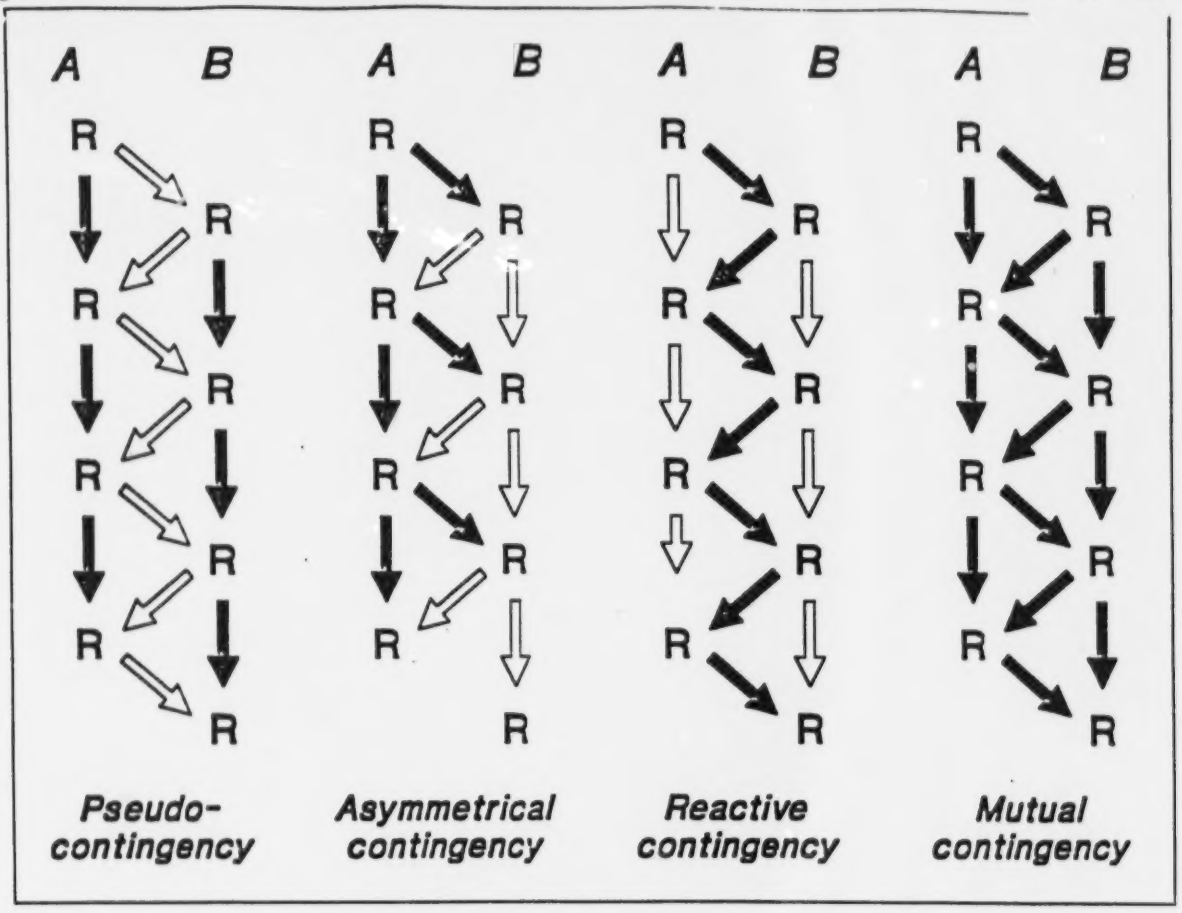

Figure 1. Styles of dyadic discourse related to contingency and orientation of speakers to internal goals. Based on Jones & Gerard (1967, p. 507).

Measures of dominance

Examiner Quantity (EQ): Quantity (jn t-units) of the examiner's talk during one task.
Candidate Quantity (CQ): Quantity (in t-units) of the candidate's talk during one task.

Measures of dominance and contingency

Examiner Topic Initiations (ET): Number of topics initiated by the examiner during one task.
Candidate Topic Initiations (CT): Number of topics initiated by the candidate during one task.

Measures of contingency

Ratification by Candidate of Examiner-initiated Topics (RET): The proportion of topics initiated by the examiner which are the topics of subsequent turns by the candidate (expressed as a percentage).

Ratification by Examiner of Candidate-initiated Topics (RCT): The proportion of topics initiated by the candidate which are the topics of subsequent turns by the examiner (expressed as a percentage).

Measures of goal orientation

Persistence of Examiner-initiated Topics (PET): Mean persistence (in t-units) of all topics initiated by the examiner during one task. The persistence of a given topic is measured by counting total t-units across both speakers from first to last mention of the topic.

Persistence of Candidate-initiated Topics (PCT): Mean persistence (in t-units) of all topics initiated by the candidate during one task.

Figure 2. Discourse variables

Table 1. Differences between examiner and candidate discourse in 30 FCE oral interviews

	Quantity of Talk by		Topic Initiations by		Percentage of Topics Initiated by Which Are Ratified		Persistence of Topics Initiated by	
	Ex.	Cand.	Ex.	Cand.	Ex.	Cand.	Ex.	Cand.
	EQ	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
N	90	90	90	90	90	87	90	87
Median	12.5	21	5	7	67%	33%	4.5	4
Mean Rank	72.7	108.3	85.7	95.3	120.1	56	98.6	79.1
Wilcoxon Z	–4.57		–1.25		–8.23		–2.52	
Prob > |Z|	.0001		.2133		.0001		.0116	
Ex.-Cand. differences	*CQ* > *EQ*		Not significant		*RET* > *RCT*		*PET* > *PCT*	

N is the number of tasks. Each interview contained three tasks, for a total of 90 observations. However, in the case of three tasks the candidate did not initiate any topics, making the values of *RCT* and *PCT* meaningless. Thus, only 87 observations are reported for these two variables. Significance level set at α=.05; df=1.

Table 2. Median scores and Kruskal-Wallis tests for interview discourse related to examiner

Examiner	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
A	12	15.5	21	7	5.5	75%	50%	4.8	4.6
B	21	9	21	4	7	71%	22%	3.5	4
C	9	13	21	4	5	50%	30%	4	3.8
D	6	16.5	17	5	8	45%	20%	3.9	3.8
E	12	18	26.5	8	8	77%	36%	5.7	3.5
F	15	12	16	4	5	75%	33%	6	4
G	6	14	23.5	6.5	8	59%	40%	3.8	2.3
H	9	10	19	6	4	83%	25%	5.3	4.7
Kruskal-Wallis χ^2		9.90	4.25	7.67	4.46	21.72	13.25	13.40	6.29
p		n.s.	n.s.	n.s.	n.s.	.0028	n.s.	n.s.	n.s.
η^2 test of association						.24			

Significance level set at α=.05; df=7. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums). Median scores are shown in Tables 2–5 rather than rank sums because they are more informative.

Table 3. Median scores and Kruskal-Wallis tests for interview discourse related to theme

Theme	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
Learning	36	14.5	21	5	6.5	66%	33%	4	4
Having a Good Time	21	11	18	6	8	75%	25%	4.5	3.6
Work	33	12	21	5	6	71%	33%	6	4.4
Kruskal-Wallis χ^2		.36	.65	1.14	3.69	2.04	.53	9.25	.50
p		n.s	n.s	n.s.	n.s.	n.s.	n.s.	.0098	n.s.
η^2 test of association								.10	

Table 4. Median scores and Kruskal-Wallis tests for interview discourse related to task

Task	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
Photographs	30	12	21	5	7	77%	40%	4.7	5
Passages Task	30	7	10	3	2.5	67%	20%	4.1	2.5
Ranking Activities	30	20.5	35.5	8	10.5	67%	32%	5	4
Kruskal-Wallis χ^2		44.12	50.68	31.48	37.70	4.34	5.62	.73	18.51
p		.0001	.0001	.0001	.0001	n.s.	n.s.	n.s	.0001
η^2 test of association		.50	.57	.35	.42				.21

Significance level set at α=.05; df=2. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).

Table 5. Median scores and Kruskal-Wallis tests for interview discourse related to gender

Gender										
Ex	Cand	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
F	F	21	16	19	6	5	82%	33%	6.2	5.5
	M	15	12	21	5	7	67%	29%	4.8	3
M	F	33	13	22	5	8	64%	33%	4	4
	M	21	11	19	4	6	67%	24%	3.5	3.7
Kruskal-Wallis χ^2			3.40	2.29	2.76	1.99	13.57	4.23	16.79	3.40
p			n.s.	n.s.	n.s.	n.s.	.0036	n.s.	.0008	n.s.
η^2 test of association							.15		.19	

Significance level set at α=.05; df=3. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).

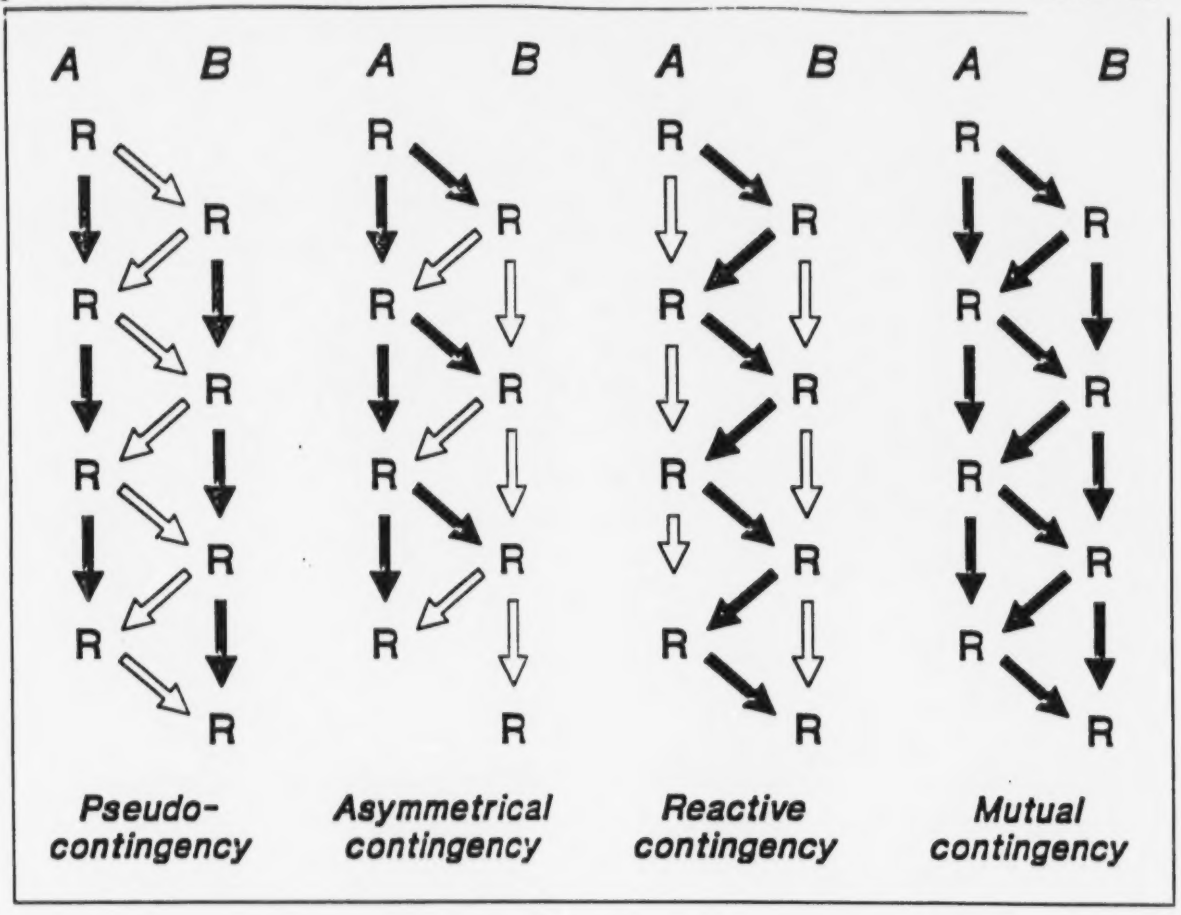

Figure 1. Styles of dyadic discourse related to contingency and orientation of speakers to internal goals. Based on Jones & Gerard (1967, p. 507).

Measures of dominance

Examiner Quantity (EQ): Quantity (jn t-units) of the examiner's talk during one task.
Candidate Quantity (CQ): Quantity (in t-units) of the candidate's talk during one task.

Measures of dominance and contingency

Examiner Topic Initiations (ET): Number of topics initiated by the examiner during one task.
Candidate Topic Initiations (CT): Number of topics initiated by the candidate during one task.

Measures of contingency

Ratification by Candidate of Examiner-initiated Topics (RET): The proportion of topics initiated by the examiner which are the topics of subsequent turns by the candidate (expressed as a percentage).

Ratification by Examiner of Candidate-initiated Topics (RCT): The proportion of topics initiated by the candidate which are the topics of subsequent turns by the examiner (expressed as a percentage).

Measures of goal orientation

Persistence of Examiner-initiated Topics (PET): Mean persistence (in t-units) of all topics initiated by the examiner during one task. The persistence of a given topic is measured by counting total t-units across both speakers from first to last mention of the topic.

Persistence of Candidate-initiated Topics (PCT): Mean persistence (in t-units) of all topics initiated by the candidate during one task.

Figure 2. Discourse variables

Table 1. Differences between examiner and candidate discourse in 30 FCE oral interviews

	Quantity of Talk by		Topic Initiations by		Percentage of Topics Initiated by Which Are Ratified		Persistence of Topics Initiated by	
	Ex.	Cand.	Ex.	Cand.	Ex.	Cand.	Ex.	Cand.
	EQ	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
N	90	90	90	90	90	87	90	87
Median	12.5	21	5	7	67%	33%	4.5	4
Mean Rank	72.7	108.3	85.7	95.3	120.1	56	98.6	79.1
Wilcoxon Z	–4.57		–1.25		–8.23		–2.52	
Prob > |Z|	.0001		.2133		.0001		.0116	
Ex.-Cand. differences	*CQ > EQ*		Not significant		*RET > RCT*		*PET > PCT*	

N is the number of tasks. Each interview contained three tasks, for a total of 90 observations. However, in the case of three tasks the candidate did not initiate any topics, making the values of *RCT* and *PCT* meaningless. Thus, only 87 observations are reported for these two variables. Significance level set at α=.05; df=1.

Table 2. Median scores and Kruskal-Wallis tests for interview discourse related to examiner

Examiner	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
A	12	15.5	21	7	5.5	75%	50%	4.8	4.6
B	21	9	21	4	7	71%	22%	3.5	4
C	9	13	21	4	5	50%	30%	4	3.8
D	6	16.5	17	5	8	45%	20%	3.9	3.8
E	12	18	26.5	8	8	77%	36%	5.7	3.5
F	15	12	16	4	5	75%	33%	6	4
G	6	14	23.5	6.5	8	59%	40%	3.8	2.3
H	9	10	19	6	4	83%	25%	5.3	4.7
Kruskal-Wallis χ^2		9.90	4.25	7.67	4.46	21.72	13.25	13.40	6.29
p		n.s.	n.s.	n.s.	n.s.	.0028	n.s.	n.s.	n.s.
η^2 test of association						.24			

Significance level set at α=.05; df=7. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums). Median scores are shown in Tables 2–5 rather than rank sums because they are more informative.

Table 3. Median scores and Kruskal-Wallis tests for interview discourse related to theme

Theme	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
Learning	36	14.5	21	5	6.5	66%	33%	4	4
Having a Good Time	21	11	18	6	8	75%	25%	4.5	3.6
Work	33	12	21	5	6	71%	33%	6	4.4
Kruskal-Wallis χ^2		.36	.65	1.14	3.69	2.04	.53	9.25	.50
p		n.s	n.s	n.s.	n.s.	n.s.	n.s.	.0098	n.s.
η^2 test of association								.10	

Table 4. Median scores and Kruskal-Wallis tests for interview discourse related to task

Task	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
Photographs	30	12	21	5	7	77%	40%	4.7	5
Passages Task	30	7	10	3	2.5	67%	20%	4.1	2.5
Ranking Activities	30	20.5	35.5	8	10.5	67%	32%	5	4
Kruskal-Wallis χ^2		44.12	50.68	31.48	37.70	4.34	5.62	.73	18.51
p		.0001	.0001	.0001	.0001	n.s.	n.s.	n.s	.0001
η^2 test of association		.50	.57	.35	.42				.21

Significance level set at α=.05; df=2. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).

Table 5. Median scores and Kruskal-Wallis tests for interview discourse related to gender

Gender										
Ex	Cand	N	*EQ*	*CQ*	*ET*	*CT*	*RET*	*RCT*	*PET*	*PCT*
F	F	21	16	19	6	5	82%	33%	6.2	5.5
	M	15	12	21	5	7	67%	29%	4.8	3
M	F	33	13	22	5	8	64%	33%	4	4
	M	21	11	19	4	6	67%	24%	3.5	3.7
Kruskal-Wallis χ^2			3.40	2.29	2.76	1.99	13.57	4.23	16.79	3.40
p			n.s.	n.s.	n.s.	n.s.	.0036	n.s.	.0008	n.s.
η^2 test of association							.15		.19	

Significance level set at α=.05; df=3. Kruskal-Wallis statistics were computed on Wilcoxon Scores (rank sums).